digested with *Nde*I and *Bam*HI, ligated into the cloning vector pET11c treated with the same pair of restriction enzymes, and transformed into *E. coli* strain NovaBlue(DE3)pLysS. Inserts in pET expression vectors were sequenced in both orientations to ensure that the plasmid constructs were free of PCR or ligation error. Nucleic acid and protein sequence analyses were carried out by Clustal method (Higgins, et al., Comput Appl Biosci, 5(2):151-153 (1989), which is hereby incorporated by reference) using MegAlign program of DNASTAR (Madison, WI).

Example 4 - Expression and Purification of *Tsp.* AK16D DNA Ligase

E. coli NovaBlue(DE3)pLysS cells containing plasmid pTAK encoding the *Tsp.* AK16D DNA ligase gene from a pET11c construct was propagated overnight at 37°C in LB medium containing 50 μg/ml ampicillin, 25 μg/ml chloramphenicol, and 0.2% glucose. Overnight cultures were diluted 100-fold into the same medium, grown until the optical density of the culture reached 0.5 at 600 nm, then induced by the addition of IPTG to a final concentration of 1 mM, and grown for an additional 4 hrs under the same conditions. Cells were collected by centrifugation, frozen/thawed at –20°C/23°C, disrupted by sonication, and clarified by centrifugation as previously described (Wetmur, et al., J Biol Chem, 269(41):25928-25935 (1994), which is hereby incorporated by reference). The resulting supernatants were heated at 70°C for 15 min to denature thermolabile *E. coli* proteins, placed on ice for 30 min to aggregate the denatured proteins, and cleared of denatured proteins by microcentrifugation for 15 min at 4°C. The partially pure DNA ligase was further purified by chromatography using 1 ml HiTrap Blue affinity column. Briefly, the column containing *Tsp.* AK16D DNA ligase was washed extensively with TE buffer (pH 7.8) containing 0.1 M NaOAc, and the ligase was eluted with TE buffer (pH 7.8) containing 2 M NaCl. After dialysis against TE buffer (pH 8.0) containing 0.2 M KCl and concentration using Centricon-30 (Amicon), protein concentration was assayed by the Bradford method with reagents supplied by Bio-Rad protein assay kit. The amount of protein was determined using BSA as the standard. The purity of the ligase was verified through 7.5% SDS (i.e. sodium dodecyl sulfate)-PAGE (i.e. polyarcylamide gel electrophoresis) analysis followed by visualizing the overloaded gel with routine Coomassie Brilliant Blue R staining.

Example 5 - Substrates And Ligation Assay

The oligonucleotide perfect match substrate was formed by annealing two short oligonucleotides (33-mer for LP3'C (SEQ. ID. No. 11) and 30-mer for Com3F (SEQ. ID. No. 12)) with a 59-mer complementary oligonucleotide (Glg). Oligonucleotides LP3'C and Glg (SEQ. ID. No. 14) were in 1.5-fold excess so that the all the 3' Fam labeled Com3F represented nicked substrates (see Luo, et al., Nucleic Acids Res, 24(15):3071-3078 (1996), which is hereby incorporated by reference). The T/G mismatch substrate was formed by annealing LP3'T (SEQ. ID. No. 13), which introduced a single base-pair mismatch at the 3'-end of the nick junction, along with Com 3'F to the complementary strand (Glg). The nicked DNA duplex substrates were formed by denaturing DNA probes at 94°C for 2 min followed by re-annealing at 65°C for 2 min in ligation buffer. The sequences of the oligonucleotides were listed below (p represents 5' phosphate group):

```
                                                    pAGTTGTCATAGTTTGATCCTCTAGTCTGGG-Fam-3' Com3
LP3'T          5'-CCCTGTTCCAGCGTCTGCGGTGTTGCGTT
LP3'C      5'-AAAACCCTGTTCCAGCGTCTGCGGTGTTGCGTC
Glg            3'-GGGACAAGGTCGCAGACGCCACAACGCAGTCAACAGTATCAAACTAGGAGATCAGACCC-5'
```

Ligation mixtures (20 µl) containing indicated amount of DNA ligase and match or mismatch substrate in the ligase buffer (20 mM Tris-HCl, pH 7.6 at room temperature; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT (i.e. dithiothreitol); 1 mM NAD^+; and 20 mg/ml BSA) were incubated at 65°C for a predetermined time. Reactions were terminated by the addition of an equal volume of stop solution (i.e. 50 mM EDTA, 80% formamide, and 1% Blue Dextran). Samples (5 µl) were electrophoresed through an 8 M urea-10% polyacrylamide GeneScan gel according to instructional manual (Perkin Elmer). The unreacted substrates were represented by the 30-mer com3F and products were represented by a ligated 63-mer in the case of the match substrate. Both the remaining substrates and ligated products were quantified using GeneScan analysis software 672 (version 2.0, Perkin Elmer).

Example 6 - Steady State Kinetics

Steady state kinetic constants were determined by measuring initial rates of the ligation reaction at a given substrate concentration (nicked DNA duplex substrate concentration ranging from 25-400 nM) and a given ligase concentration (12.5 pM for both *Tth* and *Tsp.* AK16D) in 100 µl reaction volume at 65°C. A 5 µl aliquot was removed at 0, 2, 4, 6, 8, 10 min, and mixed with 5µl of stop solution. The remaining substrate was separated from ligated product by GeneScan gel as described above. Initial rates of the ligation reactions were calculated from the generation of ligated product over time. The K_m and k_{cat} values were determined using computer software Ultrafit (Biosoft, Ferguson, MO).

Example 7 - Sequence Analysis Of Seven *Thermus* Ligase Genes

Amino acid sequence alignment of five Gram negative bacterial NAD^+-dependent DNA ligases indicates that *Tth* ligase is 93% identical to *Thermus scotoductus* ligase, 49% to *Rhodothermus marinus* ligase, 48% to *E. coli* ligase, and 38% to *Zymomonas mobilis* based on sequence data retrieved from GeneBank. Degenerate primers corresponding to highly conserved regions of these ligases were used to amplify fragments of ligase genes from seven *Thermus* strains which represent a worldwide collection: *Thermus flavus* from Japan (SEQ. ID. No. 16), *Thermus aquaticus* YT-1 (SEQ. ID. No. 15) and *Thermus sp.* AK16D from Yellowstone National Park in the United States, *Thermus filiformis* Tok4A2 (SEQ. ID. No. 17) and *Thermus filiformis* Tok6A1 (SEQ. ID. No. 18) from New Zealand, *Thermus sp.* SM32 (SEQ. ID. No. 19) from Azores, and *Thermus sp.* Vil3 (SEQ. ID. No. 20) from Portugal. The sequences of amplified ligase fragments ranging from 1.4 to 1.6 kb were determined by directly sequencing the PCR products using an ABI 373 automated sequencer. *Thermus* ligases, in general, were highly conserved during evolution as demonstrated by 85%-98% sequence identity. In contrast, the amino acid sequences of the restriction endonuclease *Taq*I and its isoschizomers from the identical strains show only 50-70% aa identities (Cao, et al., Gene, 197:205-214 (1997), which is hereby incorporated by reference). *Thermus* ligases in general show 30-40% sequence identities as compared with DNA ligases from other bacteria. The